Apr. 2011 CHINESE JOURNAL OF ENGINEERING MATHEMATICS Vol. 28 No. 2

Article ID:1005-3085(2011)02-0272-07

The Primal-dual Active Set Method for the Complementarity Problem*

LI Qing-guo, YANG Hai-jian

(College of Mathematics and Econometrics, Hunan University, Changsha 410082)

Abstract: In this paper, we deal with the convergence properties of a primal-dual active set method for the complementarity problem with T-monotone operators. We prove that the primal-dual active set method can be interpreted as a specific semismooth Newton method applied to this kind of complementarity problems. The established convergence results and numerical tests imply that the iteration number of the method is bounded by the number of the unknowns. Finally, numerical results show the efficiency of the proposed method.

Keywords: complementarity problem; the primal-dual active set method; T-monotone operator

Classification: AMS(2000) 90C33; 49M37 **CLC number:** O241.8 **Document code:** A

1 Introduction

The complementarity problem has many important applications in operation research, economic equilibrium models and in the engineering sciences[1,2]. For this reason, there is a growing interest in finding efficient and robust algorithms for solving the complementarity problem. This reflects in an increasing number of proposals of solution schemes for the complementarity problem in recent years. In these recent developments, an important role has been played by the semismooth methods, i.e., by those methods that attempt to solve the complementarity problem by first reformulating it as a semismooth system of equations and then by applying a generalized Newton method to solve this system. Another efficient way to solve complementarity problems is given by the primal-dual active set method[3]. Its basic iteration consists of two steps. In the first phase, based on a certain criterion the domain is decomposed into active and inactive parts. In the second phase, a reduced nonlinear system associated with the inactive set is solved. We remark that the above methods are based on the linear complementarity problem.

In this paper, the semismooth Newton method and the primal-dual active set method for a kind of nonlinear complementarity problems are described, respectively. It can be thought of as an extension for the linear complementarity problem, which was proposed by Hintermüller *et al*[3]. The established convergence results and numerical tests imply that the iteration number of the method is bounded by the number of the unknowns.

Received: 22 Apr 2009. **Biography**: Li Qingguo (Born in 1963), Male, Professor.

Accepted: 24 Feb 2010. Research field: applied computer science.

***Foundation item:** The National Natural Science Foundation of China (10771056); the National High Technology Research and Development Program (2006AA04A104).

2 Preliminaries

Let $N = \{1, 2, \cdots, n\}$ be an index set. For $I,\ J \subseteq N$, we denote A_{IJ} as the submatrix of the matrix $A \in \mathbf{R}^{n\times n}$ consisting of $a_{ij}\ (i \in I,\ j \in J)$, r_I as the subvector of $r \in \mathbf{R}^n$ consisting of $r_i\ (i \in I)$. Let $K \subseteq \mathbf{R}^n$, f be an operator from K to $\mathbf{R}^n$, and for all $v \in K$ be expressed by $v = v^+ + v^-$ with $v^+ = \max\{v, 0\}$, $v^- = \min\{v, 0\}$. Then the notion of T-monotone can be defined as follows.

Definition 2.1 The operator f is called T-monotone over K, if

$$\big(f(v) - f(w),\ (v - w)^+\big) \geq 0, \quad \forall\, v, w \in K, \tag{1}$$

where $(\cdot,\cdot)$ denotes the inner product. Moreover, if for all $v,\ w \in K$, $(f(v) - f(w), (v - w)^+) = 0$ is equivalent to $(v - w)^+ = 0$, then f is called strictly T-monotone over K.

In this paper, we consider the following nonlinear complementarity problem: find a solution $u^* \in \mathbf{R}^n$ such that

$$u \geq \phi, \quad f(u) \geq 0, \quad (u - \phi)^T f(u) = 0, \tag{2}$$

where $f : \mathbf{R}^n \to \mathbf{R}^n$ is a continuously differentiable, coercive and strictly T-monotone operator and $\phi \in \mathbf{R}^n$ is a given vector. It is well known that problem (2) has a unique solution[2].

Lemma 2.1 Let f be a continuous T-monotone operator over K and $I,\ J \subseteq N$ satisfying $J = N\backslash I$. For any vectors $y,\ z \in K$, if $y_I = z_I$ and $y_J \geq z_J$, then $f_I(y) \leq f_I(z)$.

Proof Let $\hat{I} = \{i \in I : f_i(y) > f_i(z)\}$ and $\hat{J} = N\backslash\hat{I}$. Without loss of generality, let $\hat{I} = \{1, 2, \cdots, k\}$ and $\hat{J} = \{k + 1, \cdots, n\}$. Suppose that $\hat{I}$ is not empty. Denote $w_1 = \{z_1 + \delta_1, \cdots, z_k + \delta_k, z_{k+1}, \cdots, z_n\}$ and $w_2 = y$, where δ_i is a positive integer for all $i \in \hat{I}$. If δ_i is small enough, then $f_i(w_1) < f_i(w_2)$ for all $i \in \hat{I}$, since the operator f is continuous. Moreover, note that $(w_1 - w_2)^+_{\hat{J}} = 0$, we have

$$0 \leq \big\langle f(w_1) - f(w_2),\ (w_1 - w_2)^+\big\rangle = \sum_{i=1}^{k} \delta_i\big(f_i(w_1) - f_i(w_2)\big) < 0,$$

which is a contradiction. Hence $\hat{I} = \emptyset$, which means that $f_I(y) \leq f_I(z)$.

Lemma 2.2 Let f be a continuous strictly T-monotone operator over K and $I,\ J \subseteq N$ satisfying $J = N\backslash I$. For any vectors $y,\ z \in K$, if $y_I \leq z_I$ and $f_J(y) \leq f_J(z)$, then $y \leq z$.

Proof Let $\hat{I} = \{i \in N : y_i \leq z_i\}$ and $\hat{J} = N\backslash\hat{I}$. Without loss of generality, let $\hat{I} = \{1, 2, \cdots, k\}$ and $\hat{J} = N\backslash\hat{I}$. Suppose that $\hat{J}$ is not empty, then by the definition of $\hat{I}$, we have that $I \subset \hat{I}$ and $\hat{J} \subset J$. Moreover

$$f_{\hat{J}}(y) \leq f_{\hat{J}}(z), \quad y_{\hat{J}} > z_{\hat{J}}, \quad y_{\hat{I}} \leq z_{\hat{I}}. \tag{3}$$

Hence $f_{\hat{J}}(z_{\hat{I}}, y_{\hat{J}}) \leq f_{\hat{J}}(y_{\hat{I}}, y_{\hat{J}}) = f_{\hat{J}}(y) \leq f_{\hat{J}}(z) = f_{\hat{J}}(z_{\hat{I}}, z_{\hat{J}})$, where the first inequality comes from (3) and Lemma 2.1. Denote $w_1 = (z_{\hat{I}}, y_{\hat{J}})$ and $w_2 = z$ by the definition of T-monotone operator, we have

$$0 \leq \big\langle f(w_1) - f(w_2),\ (w_1 - w_2)^+\big\rangle = \big\langle f_{\hat{J}}(w_1) - f_{\hat{J}}(w_2),\ (y_{\hat{J}} - z_{\hat{J}})^+\big\rangle \leq 0,$$

where the second inequality comes from $f_{\hat{J}}(w_1) \leq f_{\hat{J}}(w_2)$ and $(y_{\hat{J}} - z_{\hat{J}})^+ > 0$. Hence $\langle f_{\hat{J}}(w_1) - f_{\hat{J}}(w_2), (y_{\hat{J}} - z_{\hat{J}})^+\rangle = 0$, which means $y_{\hat{J}} - z_{\hat{J}} \leq 0$, since f is a strictly T-monotone operator. This is a contradiction to (3), and it means that $y \leq z$.

3 Semismooth Newton method

Definition 3.1[3] Let X and Y be two Banach spaces and $\mathcal{L}(X,Y)$ be denoted by the set of all linear operators on X into Y. The mapping $F: D \subset X \to Y$ is called slantly differentiable in open subset $U \subset D$ if there exists a family of mappings $G: U \to \mathcal{L}(X,Y)$ such that

$$\lim_{h\to 0} \frac{1}{\|h\|}\left\|F(x+h) - F(x) - G(x+h)h\right\| = 0, \tag{4}$$

for every $x \in U$.

We refer to G as a slanting function for F in U. Note that G is not required to be unique to be a slanting function for F in U. As shown in [4], the max-function $v \to \max\{0, v\}$ from $\mathbf{R}^n \to \mathbf{R}^n$ is slantly differentiable with a slanting function given by the diagonal matrix $G_m(v)$ with diagonal elements

$$G_m(v)_{ii} := \begin{cases} 1, & \text{if} \quad v_i > 0, \\ 0, & \text{if} \quad v_i \leq 0. \end{cases} \tag{5}$$

In the following, we restrict the Banach space to $\mathbf{R}^n$. Then the semismooth Newton method is as follows.

Algorithm 1 (semismooth Newton method)

Step 1 Given an initial vector $x^0 \in \mathbf{R}^n$, and set $k := 0$.

Step 2 Do the following sequence of vectors

$$x^{k+1} = x^k - \delta x^k, \tag{6}$$

where δx^k is the solution of the linear system

$$G(x^k)z = F(x^k), \tag{7}$$

with $G(x)$ denoting a slanting function of F at x.

Step 3 If $x^{k+1} = x^k$, then stop. Otherwise, set $k := k+1$ and return to Step 2.

In the following, we consider the equivalent form of problem (2)

$$\begin{cases} f(u) - \lambda = 0, \\ u \geq \phi, \quad \lambda \geq 0, \quad (\lambda, u - \phi) = 0, \end{cases} \tag{8}$$

Note that the complementarity system given by the second line in (8) can equivalently be expressed as

$$\mathcal{C}(u,\lambda) = 0, \quad \text{where} \quad \mathcal{C}(u,\lambda) = \lambda - \max\left\{0,\ \lambda + c(\phi - u)\right\}, \tag{9}$$

for each $c > 0$. Here the max-operation is understood to be componentwise.

Consequently, (8) is equivalent to

$$\begin{cases} f(u) - \lambda = 0, \\ \mathcal{C}(u,\lambda) = 0. \end{cases} \tag{10}$$

Define $F : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}^n \times \mathbf{R}^n$ by

$$F(u,\lambda) = \begin{pmatrix} f(u) - \lambda \\ \lambda - \max\{0, \lambda + c(\phi - u)\} \end{pmatrix}, \tag{11}$$

and note that (10) is equivalent to $F(u,\lambda) = 0$. As shown in [3], $F(u,\lambda)$ is slantly differentiable with the particular choice (5) for the slanting function of the max-function.

Let (u^*, λ^*) denote the unique solution of (8) and $x^0 = (u^0, \lambda^0)$ the initial value of the iteration. Similarly to [3], we deduce the following result. For the sake of brevity, the proof is omitted.

Theorem 3.1 Suppose that $x^* = (u^*, \lambda^*)$ is the solution to (10). Then Algorithm 1 converges superlinearly to x^*, provided that $\|x^0 - x^*\|$ is sufficiently small.

4 The primal-dual active set method

The primal-dual active set method is based on using (9) as a prediction strategy; i.e., given a current primal-dual pair (u, λ), the choice for the next inactive and active sets is given by

$$\mathcal{I} = \{i \in N : \lambda_i + c(\phi - u)_i \le 0\}, \quad \mathcal{A} = \{i \in N : \lambda_i + c(\phi - u)_i > 0\}.$$

This leads to the following primal-dual active set method for problem (2).

Algorithm 2 (the primal-dual active set method)

Step 1 Initialize u^0 and λ^0. Set $k := 0$.

Step 2 Determine the active and inactive sets by

$$\mathcal{A}_k = \{i \in N : \lambda_i^k + c(\phi - u^k)_i > 0\}, \tag{12}$$

$$\mathcal{I}_k = \{i \in N : \lambda_i^k + c(\phi - u^k)_i \le 0\}. \tag{13}$$

Step 3 Let u^{k+1} and λ^{k+1} be the solution of the nonlinear system

$$\begin{cases} f(u^{k+1}) - \lambda^{k+1} = 0, \\ u^{k+1} = \phi \quad \text{on} \quad \mathcal{A}_k, \\ \lambda^{k+1} = 0 \quad \text{on} \quad \mathcal{I}_k. \end{cases} \tag{14}$$

Step 4 Stop, or set $k = k + 1$ and return to Step 2.

We can see that the nonlinear system (14) is the realization of the following nonlinear system

$$\begin{cases} f_{\mathcal{I}_k}(u^{k+1}) = 0, \\ u_{\mathcal{A}_k}^{k+1} = \phi_{\mathcal{A}_k}, \quad \lambda_{\mathcal{I}_k}^{k+1} = 0, \\ \lambda_{\mathcal{A}_k}^{k+1} = f_{\mathcal{A}_k}(u^{k+1}). \end{cases} \tag{15}$$

Lemma 4.1 Let f be a strictly T-monotone operator. Then (15) has a unique solution.

Proof Suppose u and v are the solutions of (15). Then we have $f_I(u) = f_I(v)$, $u_J = v_J = \phi_J$ and

$$\langle f(v) - f(u), (v-u)^+ \rangle = \langle f_I(v) - f_I(u), (v_I - u_I)^+ \rangle = 0,$$

which means $(v_I - u_I)^+ = 0$, since f is a strictly T-monotone operator. That implies $v_I \le u_I$. Similarly, we have $v_I \ge u_I$. Hence $u = v$.

Theorem 4.1 Let f be a continuous strictly T-monotone operator and $u^* \in \mathbf{R}^n$ be the unique solution of problem (2). Then the sequence $\{u^k\}$ given by Algorithm 2 satisfies

(i) $u^k \le u^{k+1}$ for all $k \ge 1$ and $\phi \le u^k$ for all $k \ge 2$;

(ii) $u^k \to u^*$ at most n iterations.

Proof Firstly, we show that $u^k \le u^{k+1}$ for all $k \ge 1$. Observe that for all $k \ge 1$ the complementarity property

$$\lambda_i^k = 0 \quad \text{or} \quad u_i^k = \phi_i, \quad \forall\, i \in N, \quad k \le 1 \tag{16}$$

holds. For $i \in \mathcal{A}_k$, we have $\lambda_i^k + c(\phi - u^k)_i > 0$, and hence by (16) either $\lambda_i^k = 0$, which implies $u_i^k < \phi_i$, or $\lambda_i^k > 0$, which implies $u_i^k = \phi_i$. In particular, for $i \in \mathcal{A}_k$ we have

$$\lambda_i^k \ge 0 \quad \text{or} \quad u_i^k \le \phi_i \tag{17}$$

at each iteration $k \ge 1$. In a similar fashion, for $i \in \mathcal{I}_k$ we obtain

$$\lambda_i^k \le 0 \quad \text{or} \quad u_i^k \ge \phi_i. \tag{18}$$

Note that, for $i \in \mathcal{I}_k$, $f_i(u^{k+1}) - f_i(u^k) = \lambda_i^{k+1} - \lambda_i^k = 0 - \lambda_i^k \ge 0$, where the first equality comes from the first line in (12), and the inequality comes from (18). And $u_i^k \le \phi_i \le u_i^{k+1}$ holds for $i \in \mathcal{A}_k$. Hence, we have $u^k \le u^{k+1}$ for all $k \ge 1$ by Lemma 2.2.

Next we show that u^k is feasible for all $k \ge 2$. Due to the monotonicity of u^k it suffices to show that $u^2 \ge \phi$. Let $V = \{i \in N : u_i^1 < \phi\}$. For $i \in V$ we have $\lambda_i^1 = 0$ by (16), and hence $\lambda_i^1 + c(\phi - u^1)_i > 0$, which implies $i \in \mathcal{A}_1$. Since $u^2 = \phi$ on $\mathcal{A}_1$ and $u^2 \ge u^1$, it follows that $u^2 \ge \phi$.

Turning to the feasibility of λ^k, assume that for a pair of $(\bar{k}, i)$, $\bar{k} \ge 1$, we have $\lambda_i^{\bar{k}} < 0$. Then it is necessary $i \in \mathcal{A}_{\bar{k}-1}$, $u_i^{\bar{k}} = \phi_i$ and $\lambda_i^{\bar{k}} + c(\phi - u^{\bar{k}})_i < 0$. It follows that $i \in \mathcal{I}_{\bar{k}}$, $\lambda_i^{\bar{k}+1} = 0$ and $\lambda_i^{\bar{k}+1} + c(\phi - u^{\bar{k}_1})_i \le 0$, since $u_i^{\bar{k}+1} \ge \phi_i$, $k \ge 1$. Consequently, $i \in \mathcal{I}_{\bar{k}+1}$ and by induction $i \in \mathcal{I}_k$ for all $k \ge \bar{k} + 1$. This means that $\mathcal{I}_k \subseteq \mathcal{I}_{k+1}$.

We now prove that the method terminates in at most n iterations. Since $\mathcal{I}_k$ can not contain more than n elements, it follows from $\mathcal{I}_k \subseteq \mathcal{I}_{k+1}$ that there is an index $k_0 \in N$ such that $\mathcal{I}_{k_0} = \mathcal{I}_{k_0+1}$. Using (18) with $k = k_0 + 1$, we obtain

$$u_{\mathcal{I}_{k_0}}^{k_0+1} = u_{\mathcal{I}_{k_0+1}}^{k_0+1} \ge \phi_{k_0+1}.$$

Similarly, it follows from (17) that

$$\lambda_{\mathcal{A}_{k_0}}^{k_0+1} = \lambda_{\mathcal{A}_{k_0+1}}^{k_0+1} \ge 0.$$

Since $(u^{l_0+1} - \phi)_{\mathcal{A}_{k_0}}^{k_0+1} = 0$ and $\lambda_{\mathcal{I}_{k_0}} = 0$ in view of (14), respectively, we see that u^{k_0+1} is a solution of problem (2), which is unique from Lemma 4.1.

Remark 4.1 We can choose $u^0 = \phi$ and $\lambda^0 = f(\phi)$ as initialize iterate vectors. From the proof of Theorem 4.1, $\mathcal{I}_k = \mathcal{I}_{k+1}$ can be seen as a terminal condition of Algorithm 2. We can also conclude that Algorithm 1 based on (7) is equivalent to Algorithm 2 for the complementarity problem with T-monotone operator. For the sake of brevity, the proof is omitted.

5 Numerical experiments

Let $\Omega = (0,1)\times(0,1)$ and consider the following nonlinear complementarity problem[4]: find $u \in K$, such that

$$u \geq 0, \quad -\triangle u + f(u,x,y) \geq 0, \quad u^T\big(-\triangle u + f(u,x,y)\big) = 0, \tag{19}$$

where $K = \{u : u \geq 0\}$, $f(u,x,y) = u/(1+u) + 10x + y - 8$. We discretize (19) by using the standard five-point finite difference method on a uniform grid: $h = 1/(n+1)$, where n denotes the number of mesh nodes in $x-$ or $y-$directions ($N = n^2$ is the total number of unknowns). Then (19) reduces to the complementarity problem with the T-monotone operator[4].

We have conducted the following experiments: comparing the Algorithm 2 with the project successive overrelaxation (denoted by PSOR) method[5] and the multiplicative Schwarz (denoted by Schwarz) method[6]. In the first test, we consider the numerical solution of (19). To be precise, we use PSOR to solve (19) with $N = 400$. Numerical results are listed in Table 1. From the table, we see that the optimal factor is about $\omega = 1.8$.

Table 1: Affect of relaxation parameter ω in PSOR method

ω	iter.	cpu	ω	iter.	cpu
0.1	7111	13.015	0.3	2396	4.688
0.5	1345	2.859	0.7	868	1.984
0.9	591	1.515	1.1	408	1.171
1.3	275	0.921	1.5	172	0.703
1.7	79	0.578	1.75	54	0.515
1.8	52	0.500	1.85	57	0.500
1.9	81	0.547	1.99	670	1.656

We compare the three algorithms from the view of iteration numbers and execution times. In Schwarz, we partition $N = N_1 \cup N_2$ into two equal parts with the overlapping size $O(\frac{1}{10})$, and the corresponding subproblems are solved by PSOR with the relaxation parameter $\omega = 1.8$. In Algorithm 2, the parameter $c = 1$ and the nonlinear systems are solved by the Gauss-Seidel method. We mainly consider the affect of dimension (denoted by N) to the performance of each algorithm. In each experiment, $K = \{u \in \mathbf{R}^N : u \geq 0\}$ and $u^0 = 0$. The tolerance in the subproblems of Schwarz was chosen to be equal to 10^{-4} in the $\|\cdot\|_2$-norm, while in the outer iterative processes of Schwarz was chosen to be equal to 10^{-6} in the $\|\cdot\|_2$-norm. The tolerance in the nonlinear system of Algorithm 2 was chosen to be equal to 10^{-6} in the $\|\cdot\|_2$-norm.

Table 2 gives the history of iterations and execution times for the above iterative methods. From the view of iterations, we can easily see that Algorithm 2 is better than PSOR and Schwarz in iteration numbers. Moreover, the behavior of Schwarz is better than that of PSOR. From the view of execution times, the domain decomposition approach does not seem reasonable if there is no multiprocessor system to be used. And the behavior of Algorithm 2 is much better than any other solution methods regardless of the dimension being large or small.

Table 2: Comparison of iteration numbers and execution times (iter./cpu)

N	PSOR	Schwarz	Algorithm 2
100	34/0	18/0.046	4/0.015
400	46/0.464	33/1.471	5/0.593
900	103/4.625	52/11.359	7/5.593
1600	221/26.687	74/53.609	10/31.859
2500	368/102.140	103/225.781	12/120.125
3600	548/325.234	134/752.250	14/450.609

From the above numerical results, we can conclude that the proposed methods are effective for the complementarity problem with the T-monotone operator.

References:

[1] Ferris M C, Pang J S. Engineering and economic applications of complementarity problems[J]. SIAM Review, 1997, 39: 669-713

[2] Harker P T, Pang J S. Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications[J]. Mathematical Programming, 1990, 48: 161-220

[3] Hintermüller M, *et al.* The primal-dual active set strategy as a semismooth newton method[J]. SIAM Journal on Optimization, 2002, 13: 865-888

[4] Zeng J P, Zhou S Z. A domain decomposition method for a kind of optimization problems[J]. Journal of Computational and Applied Mathematics, 2002, 146: 127-139

[5] Glowinski R, *et al.* Numerical Analysis of Variational Inequalities[M]. Amsterdam: North-Holland, 1981

[6] Yang H J, *et al.* A multiplicative Schwarz iteration scheme for solving the linear complementarity problem with an H-matrix[J]. Linear Algebra and its Applications, 2009, 430: 1085-1098

求解互补问题的原始对偶起作用集算法

李庆国， 杨海建

(湖南大学数学与计量经济学院，长沙 410082)

摘 要: 在本文中我们得到了求解带T-单调算子的互补问题的原始对偶活跃集算法的收敛结果．当原始对偶活跃集算法求解此类互补问题时，此算法可以作为一类特殊的半光滑牛顿法．收敛结果和数值试验说明了此算法的迭代次数不超过问题未知数的个数．最终，计算结果表明此算法的可行性．

关键词: 互补问题；原始对偶活跃集算法；T-单调算子